ON A FUNCTION WHICH OCCURS IN THE LAW OF THE MEAN*

By E. R. Hedrick

1. Introduction. The formula

$$f(x+h)-f(x)=h\cdot f'(x+\theta h)=h\cdot f'(x+\xi), \qquad 0<\theta<1,$$

is known to every student of mathematics under the title "the law of the mean" or a synonym. It is usually proved that the formula is correct if $f(x)$ is defined in an interval $a \leqq x \leqq b$ and if the derivative $f'(x)$ exists in the continuum $a < x < b$.†

The quantity ξ which occurs in the formula is evidently a function of the two independent variables x and h and is defined for all values of x and h for which x and $x+h$ both lie in the interval in which $f'(x)$ exists. The purpose of the present paper is to discuss the properties of this quantity ξ, and we shall write

$$\xi = \xi\,(x, h)$$

whenever it is desired to emphasize its functional character.

In studying the function $\xi(x, h)$ it occurs to one immediately that if $\xi(x, h)$ is continuous, the derivative $f'(x)$ is surely continuous (see p. 183). A slight inspection would tend to convince one that the converse is true—at least in a limited sense—i. e., that if $f'(x)$ is continuous, $\xi(x, h)$ is a continuous function of h when x is constant. This conclusion is, however, erroneous. It will be shown that $\xi(x, h)$ may be utterly discontinuous, in the sense that ξ fails to take on values which lie between values which it does take on, even when $f'(x)$ exists and is continuous everywhere (see p. 190).

The fallacies which depend upon the utter discontinuity of ξ are known, but not well-known. While pointing out their dangers, it may be well to remark that certain proofs may be conducted without error along the same general lines as these fallacious arguments, if the function ξ be fully understood.

Incidentally it will become evident that the law of the mean may hold in its present form in case the derivative $f'(x)$ does not even exist at certain points, even when we permit x and $x+h$ to assume any values whatever; and generalized statements of the law hold even in more tenuous cases.

---

* This paper is based partially upon a paper entitled "On the Law of the Mean," read by the writer at the summer meeting, 1903, of the American Mathematical Society, and partially on another paper entitled "The function $\xi(h)$ in the Law of the Mean," read by the writer at the April meeting, 1906, of the Chicago Section of the Society.

† A proof of this fact will incidentally occur as a corollary to a theorem of this paper.

Finally, certain unusual examples in the theory of maxima and minima which are suggested by the paper will be presented and commented upon for their useful warning against too great assumption in that theory.

2. Limits over an assemblage. *A function $f(x)$ of a real variable x is said to approach a limit K as x approaches a particular value a if, corresponding to any pre-assigned positive number ${}^1\epsilon$, another positive number ${}^2\delta$ exists, such that*

$$(1) \qquad |f(x) - K| < {}^1\epsilon \quad \text{whenever} \quad 0 < |x - a| < {}^2\delta$$

*where x may be any number whatever which satisfies the last inequality.** I shall say for distinctness and emphasis that $f(x)$ approaches K *with respect to the continuum* or *in the ordinary sense* whenever these conditions are fulfilled. The left-handed superscripts indicate the order of choice of ϵ and δ.

It may happen that the state of affairs described by (1) holds only in case x is allowed to assume the values which belong to a certain assemblage (E). In this case I shall say that $f(x)$ approaches K *with respect to the assemblage* (E) or simply *over* (E), and I shall write

$$\lim_{x=a} {}_{(E)} f(x) = K.$$

It is evidently trivial to consider a case in which a is not a limiting point of (E), since any function evidently approaches any desired constant for such an assemblage. We shall therefore assume that $x = a$ is a limiting point of (E) unless the contrary is stated.

In particular, if (E) consists of all points to the left (right) of a, and if the preceding inequalities are satisfied when x takes on any value in (E), *then $f(x)$ will be said to approach K from the left (right).*

If $f(a) = K$ and if $f(x)$ approaches K with respect to an assemblage (E) as x approaches a, we shall naturally say that $f(x)$ *is continuous at a with respect to* (E) or simply *over* (E).

It is now very easy to write down the following theorems, at least some of which are important:—

* It should be noted, especially for what follows, that this definition absolutely eliminates any consideration of *time*; nothing whatever is said concerning the order in which the values of x are taken on, and all notion of succession is eliminated from the vulgar notion "approach."

I. *Let* (E) *be an assemblage with a single limiting point* a. *Then* (E) *is countable and its members may be written in a sequence*

$$(E): \quad x_1, x_2, x_3, \cdots, x_n, \cdots;$$

if $f(x)$ *approaches* K *over* (E) *we shall have* $\lim_{i=\infty} f(x_i) = K$. For ${}^2N$ can be chosen so large that $|x_n - a| < {}^1\epsilon$ whenever $n \geqq {}^2N$.

II. *If* $f(x)$ *approaches a limit* K *with respect to any assemblage* (E), *it approaches* K *with respect to any sub-assemblage of* (E).

III. *If* $f(x)$ *approaches* K *over* (E), *the assemblage* (E) *may be increased by adjoining the points of any assemblage for which* $x = a$ *is not a limiting point.*

IV. *More generally, if* $f(x)$ *approaches* K *over each of a finite number of assemblages* $(E)_1, (E)_2, \cdots, (E)_n$, *the limit is the same over the assemblage* (E) *formed by uniting all the* $(E)_i$. For, given ϵ, a δ can be found for each $(E)_i$; the least of these δ's may be used for (E).

V. *In particular, if the finite set of assemblages* $(E)_i, (i = 1, 2, \cdots n)$, *exhaust all possible numbers in the neighborhood of* $x = a$, (E) *is the continuum, and* $f(x)$ *is continuous in the ordinary sense.*

VI. *On the other hand, if* $f(x)$ *approaches* K *with respect to any assemblage* (E), *any number whatever* $b \neq a$ *is a member of some assemblage* (F) *for which* $f(x)$ *approaches* K. For at least one such (F) is formed by adjoining b to (E).

VII. *It follows that even though* $f(x)$ *approaches* K *for each of a set of assemblages* (E) *which exhaust all possible numbers,* $f(x)$ *does* NOT *necessarily approach* K *in the ordinary sense.* This follows directly from VI and from the knowledge of a single example of a function which approaches no limit in the ordinary sense, but which does approach a limit over some one assemblage. Such an example is $y = f(x) = \sin(1/x)$; this function approaches no limit as x approaches zero, in the ordinary sense, but approaches the limit zero over the assemblage $(E) = 1/n\pi$ $(n = 1, 2, \cdots)$. Hence, by VI, $f(x)$ approaches zero over each of a set of assemblages which exhaust all possible numbers. Even if the set of (E)'s for which $f(x)$ approaches K and which exhaust all possible numbers is known to be a countable (or enumerable) infinite set, we cannot conclude that $f(x)$ approaches K in the ordinary sense. [The example $f(x) = 0$ when x is irrational, $f(x) = i \cdot x$ when $x = \dfrac{p}{q^i}$, where p is prime to q^i and q is not a perfect power, p, q, i being all positive integers,

is an example of this kind. For $\lim f(p/q^i) = 0$ for any fixed value of i, and also $\lim f(x) = 0$ if x is irrational, but $f(x)$ does not approach zero in the ordinary sense, as is seen by actually drawing the figure.*] *If the approach is uniform, however, the conclusion that $f(x)$ approaches K is correct.*

VIII. There is another theorem which is applicable in cases not covered by V. For *if we know that $f(x)$ approaches K for every possible assemblage, or even if we know merely that $f(x)$ approaches K for every possible sequence, we may conclude that $f(x)$ approaches K in the ordinary sense.*† For, if $f(x)$ does not approach K, there exists a positive number ϵ for which relations (1) are false for every δ. Choosing $\delta_i < 1/2^i$, we shall then have $|f(x_i) - K| > \epsilon$ for some one $x_i < \delta_i$. Then $f(x)$ surely does not approach K over the sequence of the x_i's.

These results evidently apply also in testing a function for *continuity over an assemblage* (E) at a point $x = a$, provided that $f(a) = K$.

Again, we may substitute the notion of *oscillation* for that of approaching a limit. If we consider all the points of an assemblage (E) which lie in a variable interval about $x = a$, it is evident that there is one and only one lower limit of the upper limits of $f(x)$ on (E) in all such intervals. Let us call this L. Likewise there is only one upper limit of the lower limits of $f(x)$ on (E) in all such intervals, which we call l. Then *the oscillation at $x = a$ with respect to (E) is $\omega_{(E)} = L - l$*, and it is evidently necessary and sufficient that $\omega_{(E)} = 0$ in order that $f(x)$ be continuous at $x = a$ with respect to (E). This notion may be generalized essentially;‡ and theorems analogous to those above may be written down.

3. The function $\xi(x, h)$. If $f(x)$ be defined in the interval $a \leqq x \leqq b$, the quotient

$$Q(x, h) = \frac{f(x + h) - f(x)}{h}$$

is a function of x and h defined for values of x and h for which x and $x + h$ both lie in the interval $a \leqq x \leqq b$. If this function $Q(x, h)$, for a constant value of x, approaches a limit K(as h approaches zero) over any assemblage

---

* See also *Bull. Amer. Math. Soc.*, vol. 11, footnote, p. 321, March, 1905.

† A generalization of this theorem is obvious: If $f(x)$ approaches K for every possible sequence which can be formed from the members of any assemblage (E), then $f(x)$ approaches K over (E).

‡ I shall do this in another paper.

(E) for which $h = 0$ is a limiting point, K *is said to be the derivative of* $f(x)$ *with respect to* (E), or simply *over* (E). If (E) is the continuum, the derivative $f(x)$ exists *over the continuum*, or simply *in the ordinary sense*. Otherwise we shall write

$$K = D_{(E)} f(x).$$

Let us consider the auxiliary function

$$\phi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Let us suppose that $\phi(x)$ is at a maximum (or at a minimum) at some point $x = c$, $a < c < b$. This will certainly be the case, for example, if $f(x)$ is continuous, since $\phi(b) = \phi(a) = 0$. Then if $D_{(E)}\phi(c)$ exists, where (E) denotes any essentially two-sided assemblage,* we must have

$$D_{(E)}\phi(c) = D_{(E)} f(x) - \frac{f(b) - f(a)}{b - a} = 0.$$

IX. *Hence if* $\phi(x)$ *is at a maximum (or at a minimum) at* $x = c$, $a < c < b$ *(in particular, if* $f(x)$ *is continuous) and if* $D_{(E)} f(c)$ *exists for a two-sided assemblage* (E), *we shall have*

$$f(b) - f(a) = (b - a) D_{(E)} f(c), \qquad a < c < b.$$

This is the law of the mean. The requirement that $f(x)$ be continuous in the interval $a \leqq x \leqq b$ and that the ordinary derivative $f'(x)$ exist for $a < x < b$ is evidently *sufficient*, but is *in no sense necessary* for this result.

Replacing a by x and $b - a$ by h, the formula may be written in the form

$$Q(x, h) = \frac{f(x + h) - f(x)}{h} = f'(x + \xi), \qquad 0 < \xi/h < 1,$$

at least if $f'(x)$ exists throughout an interval in which x and $x + h$ lie. That this formula holds in a very general manner in the precise form just given, even when $f(x)$ is discontinuous between x and $x + h$, is evident from such an example as $y = f(x) = \sin(1/x)$, which is essentially discontinuous at $x = 0$. If we add $f(0) = 0$, the function $f(x)$ is defined for every value of x, and it is easy to see that the preceding formula holds when x and $x + h$ assume absolutely any values whatever.

* That is, an assemblage in which there are points on both sides of $x = a$ in any neighborhood of $x = a$.

We shall now assume, however, that $f'(x)$ exists in the ordinary sense. The number ξ depends upon the choice of both x and h, and may have several values, or even an infinite number, for one pair of values of x and h. If x is fixed, ξ has at least one value for every value of h and that value is numerically less than h; i. e., in a plane whose coordinates are ξ and h, ξ always lies between the h-axis and the line $\xi = h$.

There is a common fallacy which we may revise into an accurate proof. Since $f'(x)$ exists, we shall have, at any point $x = k$,

$$\lim_{h=0} Q(k, h) = f'(k)$$

and since

$$Q(k, h) = f'(k + \xi), \qquad 0 < \xi/h < 1$$

we shall also have

$$\lim_{h=0} f'(x + \xi) = f'(k).$$

Since ξ is numerically less than h, $\xi(k, h)$ approaches zero with h; and we might be led to infer that

$$\lim_{\xi=0} f'(k + \xi) = f'(k),$$

i. e., that $f'(x)$ is continuous at $x = k$. *This is false*, however; and the fallacious nature of the supposed conclusion is exposed by a single well-known example in which $f'(x)$ exists and is discontinuous:

$$\begin{aligned} y &= f(x) = x^2 \sin(1/x^2) \qquad && \text{if } x \neq 0, \\ y &= f(x) = 0 && \text{if } x = 0. \end{aligned}$$

The function $f(x)$ thus defined is evidently continuous, and its derivative is

$$\begin{aligned} y' &= f'(x) = 2x \sin(1/x^2) - (2/x)\cos(1/x^2) \qquad && \text{if } x \neq 0, \\ y' &= f'(x) = \lim_{h=0} h \sin(1/h^2) = 0 && \text{if } x = 0. \end{aligned}$$

These forms show that $f'(x)$ exists for every value of x, and that $f'(x)$ is discontinuous at $x = 0$, for $f'(x)$ exceeds any fixed number in any interval about $x = 0$. This example illustrates also the correct conclusions which follow.

In fact the function $\xi(0, h)$ does not assume all values near zero for values of h which lie in *any* limited interval whatever. For $|f'(x)|$ exceeds any fixed constant ${}^2C$ inside of any interval $(-{}^1\epsilon, 0)$ and also inside $({}^1\epsilon, 0)$. Such is the real nature of the fallacy in any case; for *if* $\xi(k, h)$ *is a continuous function of* h, or in fact even *if having assigned* ${}^1\rho$ *we can find* ${}^2\lambda$ *such that* $|\xi|$

takes on all values less than λ *for some values of* $|h|$ *less than* ρ, *the preceeding argument really shows that* $f'(x)$ is *continuous at* $x = k$. For, having chosen $\epsilon_1, \epsilon_2, \cdots, \epsilon_n, \cdots$ $(\lim_{i=\infty} \epsilon_i = 0)$ we can choose $\rho_1, \rho_2, \cdots, \rho_n, \cdots$ $(\lim_{i=\infty} \rho_i = 0)$ in such a way that

$$|Q(k, h) - f'(k)| < \epsilon_i, \text{ whenever } |h| < \rho_i.$$

If, corresponding to each ρ_i a quantity λ_i can be found such that $|\xi|$ takes on every positive value less than λ_i for some value of $|h|$ less than ρ_i we shall have

$$|Q(k, h) - f'(k)| \equiv |f'(k + \xi) - f'(k)| < \epsilon_i \text{ whenever } |\xi| < \lambda_i,$$

which expresses that $f'(x)$ is continuous at $x = k$.

X. *A sufficient condition that* $f'(x)$ *should be continuous is that corresponding to any positive number* [1]ρ *another positive number* [2]λ *exists, such that* $|\xi|$ *takes on all values less than* λ *for some value of* $|h|$ *less than* ρ, *for both positive and negative values of* ξ.

Geometrically, a sufficient condition that a variable tangent should approach a fixed tangent is that in any interval about the fixed point of tangency a subinterval exists such that any tangent in the subinterval is parallel to some secant in the original interval through the fixed point.

Whether this condition be satisfied or not, we can always find *an assemblage* (E) *with respect to which* $f'(x)$ is continuous at $x = k$. For, having chosen h_1, there exists at least one corresponding ξ different from zero, say ξ_1. Setting $h_2 = \xi_1$, we can find a corresponding $\xi_2 \neq 0$. Proceeding in the same manner we find ξ_i, $i = 1, 2, \cdots$, $|\xi_{i+1}| < |\xi_i|$, where $\lim_{i=\infty} \xi_i = 0$. Since $Q(k, h)$ approaches $f'(k)$ as h approaches zero, $\lim_{i=\infty} Q(k, h_i) = f'(k)$, and therefore

$$\lim_{i=\infty} f'(k + \xi_i) = \lim_{i=\infty} Q(k, h_i) = f'(k),$$

which shows that $f'(x)$ approaches $f'(k)$ over $(E) = k + \xi_1, k + \xi_2, \cdots$ as x approaches k. Hence we may say

XI. *A derivative which exists at every point in an interval is continuous at each point of the interval with respect to at least one essentially two-sided assemblage.*

XII. *It follows that if the derivative* $f'(x)$, *which is supposed to exist at every point, approaches any limit whatever in the ordinary sense, as* x *approaches* k, *that limit is* $f'(k)$.

XIII. *It is evident from* VI *that if $f'(x)$ exists near $x = k$ the assemblages with respect to which $f'(x)$ approaches $f'(k)$ as x approaches k exhaust all possible numbers.* But there are still other assemblages for which $f'(x)$ does not approach $f'(k)$ unless $f'(x)$ is continuous at $x = k$. (See VII).

It is easy to generalize these statements somewhat. For the derivative $f'(x)$, which is supposed to exist, surely takes on any value C $(A < C < B)$ at some point $x = t$ $(a < t < b)$ if it takes on the values A and B at $x = a$ and $x = b$, respectively. For, $Q(x, h)$ being continuous in both x and h, $h_0 > 0$ can surely be choosen so small that $Q(a, h_0) < C$ and $Q(b - h_0, h_0) > C$. Hence $Q(c, h_0) = C$, since Q is continuous in x. Then, by the law of the mean $Q(c, h_0) = f'(c + \xi) = C$, where $s = c + \xi$ lies between a and b since $h_0 > \xi > 0$ and $a < c < b - h_0$. This fact is well known, i. e., *in passing from one value to another, any derivative $f'(x)$ takes on every intermediate value provided it exists at every point.*

XIV. *It follows that if $f'(x)$ approaches two different limits m and n (as x approaches k) with respect to two different assemblages (M) and (N), then $f'(x)$ approaches any intermediate value p, $(m < p < n)$, with respect to some third assemblage (P).* For $f'(x)$ assumes values which differ from m (and also values which differ from n) by less than any preassigned number ϵ near $x = k$; hence $f'(x)$ actually assumes the value p at an infinite number of points on both sides of $x = k$, which proves the theorem. In particular, *if $f'(x)$ approaches any other value $V \neq f'(k)$ over any assemblage, then $f'(x)$ also approaches any number between V and $f'(k)$ over some assemblage.* The property mentioned in the theorem is another property of continuous functions which is shared by any existing derivative, for if $\phi(x)$ is continuous except at $x = k$, $\phi(x)$ surely has the property in question. We shall resume this discussion presently.

Again, using the previous notation (p. 183) let ξ_i and ξ_{i+1} be the values of ξ which correspond to h_i and h_{i+1}, respectively, and let $h_i = \xi_{i+1}$. Then we shall have $Q(k, h_i) = f'(k + \xi_i) = L_i$, say, for each value of i. But since $Q(k, h)$ is continuous in h, it assumes every value between L_i and L_{i+1} for some value of h between h_i and h_{i+1}. Hence the assemblage of values of h between h_i and h_{i+1} for which $Q(k, h)$ lies between L_i and L_{i+1} *has the power of the continuum.* Since $f'(x)$ also assumes all intermediate values in passing from one value to another, $f'(k + \xi)$ also assumes every values between L_i and L_{i+1} for some value of h between h_i and h_{i+1}. Hence for the values of h found above we shall have $Q(k, h) = f'(k + \xi)$ where h lies between h_i and h_{i+1} and

where ξ lies between ξ_i and ξ_{i+1}. Thus the assemblage of values of ξ which lie between ξ_i and ξ_{i+1} and which correspond to values of h between h_i and h_{i+1} is an assemblage *which has the power of the continuum.* Let us then consider the sequence $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ and also the interpolated values of ξ just found. The total assemblage (T) formed by uniting all these has the power of the continuum in any interval about $\xi = 0$. On the other hand it is evident that $f'(k + t)$ approaches $f'(k)$ with respect to this assemblage (T). It follows that

XV. *If $f'(x)$ exists in any interval, it is continuous at every point of that interval with respect to a certain assemblage (T) which has the power of the continuum on each side of the point in any interval whatever about the point.*

It would be fallacious to conclude that $f'(x)$ is continuous at any point with respect to the assemblage of values which $k + \xi$ may assume, and this conclusion might lead to as grave error as the previous fallacy. For example, in the previous exercise $y = f(x) = x^2 \sin(1/x^2)$, ξ may actually take on any value whatever for some value of h. Hence the conclusion that $f'(x)$ is continuous for the assemblage of all possible $k + \xi$ would lead to the erroneous conclusion that $f'(x)$ is continuous in the ordinary sense.

4. Derivatives as limits of continuous functions. The function $Q(x, h)$ approaches $f'(x)$, if $f'(x)$ exists, for any sequence of values of h whose limit is zero. Selecting such a sequence:

$$1) \qquad h_1, h_2, \cdots, h_n, \cdots$$

where the h's do not depend on x,* the corresponding values of Q, viz:

$$2) \qquad Q(x, h_1), Q(x, h_2), \cdots, Q(x, h_n), \cdots$$

form a sequence of continuous functions of x, and the series

$$3) \qquad Q(x, h_1) + [Q(x, h_2) - Q(x, h_1)] + [Q(x, h_3) - Q(x, h_2)] + \cdots + [Q(x, h_{i+1}) - Q(x, h_i)] + \cdots$$

is a series of continuous functions which surely converges to the sum $f'(x)$ for any value of x for which $f'(x)$ exists. The results concerning such series are therefore immediately applicable. If the series (3) converges uniformly, for example, the derivative $f'(x)$ is continuous, for the sum of a uniformly con-

* It is easy to see that most of what follows is independent of this assumption.

vergent series of continuous functions is continuous. Such is the case, in particular, if

$$|f'(x) - Q(x, h_i)| < {}^1K \qquad \text{whenever } i > {}^2N;$$

or again if

$$|Q(x, h_n) - Q(x, h_p)| < {}^1K \qquad \text{whenever } a > n > {}^2N,$$

where N is independent of x. That is to say, if $Q(x, h)$ approaches its limit $f'(x)$ *uniformly* for any constant sequence 1) and for all values of x, then $f'(x)$ is continuous. Any example of a derivative which exists and is discontinuous (e. g., the example given above) is also an example of a series of continuous functions whose limit exists and is discontinuous, i. e., an example of non-uniform approach; and that for any sequence of h's of the type 1).

On the other hand, suppose that $f'(x)$ is continuous. Then since any continuous function is uniformly continuous, $f'(x + t)$ approaches $f'(x)$ uniformly. Therefore

$$|f'(x) - Q(x, h_i)| < {}^1K \text{ whenever } i \geqq {}^2N, \text{ independent of } x,$$

for $Q(x, h_i) = f'(x + \xi_i)$ and $f'(x + \xi_i)$ surely approaches $f'(x)$ uniformly. It follows that if $f'(x)$ is continuous the sequence 2) is uniformly convergent; hence the series 3) [and also the sequence 2)] has the following unusual property:

XVI. *The sum $f'(x)$ of the series* 3) *of continuous functions is continuous in an interval when and only when the series* 3) *converges uniformly in that interval.*

It is shown by Baire* that the limit of any sequence of continuous functions must be continuous at least once (and therefore an infinite number of times) in any interval in which the limit exists, and the obvious conclusion is drawn that any derivative which exists throughout an interval is continuous at at least one point in every subinterval.

If the derivative $f'(x)$ is discontinuous at any poiut $x = a$, it must be possible to assign a positive number ϵ such that $|f'(x) - f'(a)| > \epsilon$ for some x in any neighborhood of $x = a$. But since $f'(x)$ takes on all intermediate values between any two which it does take on, it follows that at any point of discontinuity $x = a$ a positive number ϵ depending upon a can be assigned, such that $|f'(x)|$ takes on every value between $|f'(a)|$ and $|f'(a)| + \epsilon$ in any neighborhood of $x = a$.

---

* Baire, Thesis: *Sur les fonctions discontinues*, p. 30 and p. 108.

XVII. *The range of values all of which a derivative actually takes on in any neighborhood of any point $x = k$ is zero only when the derivative is continuous at $x = k$.*

Let us inquire whether a derivative may be discontinuous at points which are everywhere dense in an interval throughout which it exists. Any discontinuous derivative (e. g., the example given above) is an example of the remarkable class of functions studied by Darboux, which are not continuous, but which do assume all intermediate values. It is even possible to construct a function which is discontinuous at every point of an interval but which assumes all intermediate values in passing between any two which it does assume.* Such a function is not the derivative of any function whatever, on account of Baire's theorem.

On the other hand, it is easy to write down a function which is continuous at points everywhere dense and also discontinuous at points everywhere dense; e. g., the well known function

$$F(x) = 0 \quad \text{if } x \text{ is irrational}; \; F(x) = 1/q \quad \text{if } 0 < x = p/q < 1$$

is defined whenever $0 < x < 1$, is continuous at every irrational point and discontinuous at every rational point in that interval. This function is not the derivative of any function whatever, since it does not assume all intermediate values in the necessary manner.

The question remains then whether it is possible to construct a function which has all these peculiar properties, i. e., a function such that the range of values all of which it actually takes on in the neighhorhood of a point is zero at some point and different from zero at some other point in every interval. At first sight this seems impossible, but it becomes evident that such functions actually exist if we first write the theorem:

---

* See Lebesgue, *Leçons sur l'integration* (1905), p. 90. Lebesgue defines $\phi(x)$ for $0 < x < 1$ as follows: having written x as an ordinary decimal $x = \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3} + \dots$ he considers the sequence $a_1, a_3, a_5, a_7, \dots$; if this sequence is *not* periodic he defines $\phi(x) = 0$; if the same sequence is periodic, if a_{2n-1} is the first member of a period, he defines

$$\phi(x) = \frac{a_{2n}}{10} + \frac{a_{2n+2}}{10^2} + \frac{a_{2n+4}}{10^3} + \dots$$

Then $\phi(x)$ takes on every value between 0 and 1 in every subinterval of the interval $0 \leqq x \leqq 1$; but $\phi(x)$ is discontinuous at every point.

XVIII. *Either* (a) *the derivative of any differentiable function is continuous throughout some subinterval of any interval in which it exists;* or else (b) *the derivative has the unusual property just mentioned.*

The discovery of a derivative which exists at every point and which is discontinuous at points everywhere dense in any interval would therefore lead to a remarkable example of a function. It is obvious that such a function results if we apply Hankel's principle of condensation of singularities to the preceding example $x^2 \sin(1/x^2)$, for we can write down at once a uniformly convergent series which has a derivative everywhere, which itself is discontinuous at points everywhere dense.* *Thus there actually exists functions (e. g.,*

---

* Such examples are known. The intention of this remark is to call attention to the consequence mentioned above.

Lebesgue indicates (l. c., p. 94), but does not actually write out, the following example:

$$\text{Let} \qquad \left.\begin{aligned} \phi(x, a) &= (x-a)^2 \sin \frac{1}{x-a}, \quad x \neq a \\ &= 0, \qquad\qquad\qquad\qquad x = a \end{aligned}\right\} \begin{aligned} 0 &\leqq a \leqq 1 \\ 0 &\leqq x \leqq 1 \end{aligned}$$

and let

$\phi_i(x) = \phi(x, a_i)$ where $a_1, a_2, \ldots, a_n, \ldots$ are the elements of any countable dense assemblage (E) (e. g., all rational numbers) arranged in a sequence.

Then the series

$$(1) \qquad \phi(x) = \phi_1(x) + \frac{1}{2!}\phi_2(x) + \ldots + \frac{1}{n!}\phi_n(x) + \ldots, \quad 0 \leqq x \leqq 1$$

is an absolutely and uniformly convergent series, $0 \leqq x \leqq 1$.

$$\text{Let} \quad \left.\begin{aligned} \psi(x, a) = \frac{\partial \phi(x, a)}{\partial x} &= 2(x-a) \sin \frac{1}{x-a} - \cos \frac{1}{x-a}, \quad x \neq a \\ &= 0, \qquad\qquad\qquad\qquad\qquad\qquad x = a \end{aligned}\right\} \begin{aligned} 0 &\leqq a \leqq 1 \\ 0 &\leqq x \leqq 1 \end{aligned}$$

and let $\psi_i(x) = \psi(x, a_i)$. Then the series of discontinuous functions:

$$(2) \qquad \psi(x) = \psi_1(x) + \frac{1}{2!}\psi_2(x) + \ldots + \frac{1}{n!}\psi_n(x) + \ldots$$

converges absolutely and uniformly, $0 \leqq x \leqq 1$, and the individual terms, $\psi_i(x)$, are the derivatives of the individual terms of (1). *It follows that* $\psi(x)$ *is the derivative of* $\phi(x)$, *and that* $\psi(x)$ *is discontinuous at each point of* (E).

In order to prove this we need to show first that (2) for every value of x is a uniformly convergent series of continuous functions plus at most a single discontinuous function; hence $\psi(x)$ is continuous except at points of (E) and is discontinuous at any point a_i of (E) with the same discontinuity as $\frac{1}{i!}\psi(x, a_i)$.

We must also show that $d\phi(x)/dx = \psi(x)$. *This follows from the fact that* (2) *is a uniformly convergent series of derivatives of the terms of* (1) although $\int \psi(x)dx$ does not exist in Riemann's sense.

$\psi(x)$ in footnote) which are (1) *discontinuous at some point in every interval,* (2) *continuous at some (other) point in every interval, and which* (3) *actually take on every intermediate value between any two values which they do assume.*

5. The Derivative Continuous. Let us now consider a function $f(x)$ whose derivative exists and is continuous throughout some interval $a \leqq x \leqq b$. We know that a sufficient condition for the continuity of $f'(x)$ is that $\xi(x, h)$ be a continuous function of h when x is constant, or even that $|\xi|$ assume all values less than ${}^2\lambda$ for some value of $|h|$ less than ${}^1\rho$. Conversely, *can we infer that if $f'(x)$ is continuous, it will follow that $\xi(x, h)$ is a continuous function of h when x is constant? Or even that $|\xi|$ assumes all values less than ${}^2\lambda$ for some value of $|h|$ less than ${}^1\rho$?* Such a conclusion seems to be obvious at first sight, but it is untrue.

Let us consider, for example, the function

$$\begin{aligned} y &= f(x) = x^3\left(1 + \sin\frac{1}{x}\right) && \text{if } x \neq 0, \\ y &= f(x) = 0 && \text{if } x = 0. \end{aligned}$$

The derivative is

$$\begin{aligned} y' &= f'(x) = 3x^2\left(1 + \sin\frac{1}{x}\right) - x\cos\frac{1}{x} && \text{if } x \neq 0, \\ y' &= f'(x) = 0 && \text{if } x = 0, \end{aligned}$$

which is continuous for all values of x. Nevertheless there are values of $\xi(0, h)$ which are not taken on for any value of x whatever and which lie as close to zero as we pleasė. For it is obvious that

$$f'(x) = 0 \quad \text{when } x = \frac{2}{(4n+1)\pi}$$

and that

$$f(x) = 0 \quad \text{when } x = \frac{2}{(4n+1)\pi}; \quad f(x) > 0 \quad \text{when } x \neq \frac{2}{(4n+1)\pi}.$$

Now $f'(x) \not\equiv 0$, since $f'(x) \not\equiv$ const. Moreover $f'(x) \geqq 0$ is surely false, since $f(0)$ is not a minimum to the right. Hence $f'(x) < 0$ for some value of x, say x_n, which lies between $\dfrac{2}{(4n+1)\pi}$ and $\dfrac{2}{(4n+5)\pi}$. It follows that ξ never takes on the value x_n, since in the formula

$$f(h) - f(0) = f'(\xi), \qquad 0 < \xi < h$$

the right-hand side is never negative. We may therefore conclude that

XIX. *It is entirely possible that* $f'(x)$ *exists and is continuous throughout an interval about a point* $x = k$, *and yet that the assemblage of values which the function* $\xi(k, h)$ *never assumes has the power of the continuum.* For since $f'(x)$ is continuous, it assumes negative values (say) for all values of x in the neighborhood of any point where it is negative.

6. Maxima and Minima. Some of the remarks made above and several of the examples are suggestive in the theory of ordinary maxima and minima. It follows directly from the law of the mean that *if*

$$f'(a) = 0, \textit{ and } f'(a+h) < 0, \textit{ and } f'(x-h) > 0, \textit{ whenever } 0 < h < \delta$$

then $f(x)$ *is at a proper maximum * for* $x = a$. It seems at first sight obvious that the converse would also be true and such a statement can be found without extensive search. Let us examine a few problems.

In the example

$$y = f(x) = x^2\left(1 + \sin\frac{1}{x^2}\right) + e^{-\frac{1}{x^2}} \qquad \text{if } x \neq 0,$$

$$y = f(x) = 0 \qquad \text{if } x = 0,$$

which is a slight modification of that on p. 182, the derivative is

$$y' = f'(x) = 2x\sin\frac{1}{x^2} - \frac{2}{x}\cos\frac{1}{x^2} + 2x + \frac{2}{x^3}e^{-\frac{1}{x^2}} \qquad \text{if } x \neq 0,$$

$$y' = f'(x) = 0 \qquad \text{if } x = 0.$$

Since the term $(2/x)\cos(1/x^2)$ determines the sign of $f'(x)$ for values of x for which $|\cos(1/x^2)|$ is sufficiently large near $x = 0$, it is evident that $f'(x)$ takes on both positive and negative values on each side of $x = 0$ inside of any interval about $x = 0$. It follows that the preceding method leads to no result. Moreover $f''(x)$ does not even exist for $x = 0$. Hence the only usual method for determining whether $x = 0$ gives a minimum is by examining the function directly. It is evident, however, that $f(x)$ is at a proper minimum at $x = 0$.

An objection might be made that $f'(x)$ is discontinuous at $x = 0$ in the preceding example. To meet such reasonable objection, we may cite the example

$$y = f(x) = x^4\left(1 + \sin\frac{1}{x}\right) + e^{-\frac{1}{x^2}} \qquad \text{if } x \neq 0,$$

$$y = f(x) = 0 \qquad \text{if } x = 0,$$

* That is $f(a) < f(a \pm h)$, $0 < h < \delta$. If merely $f(a) \leqq f(a \pm h)$ the minimum is said to be *improper*.

which is a slight modification of the example of p. 189. Here the derivative is

$$y' = f'(x) = 4x^3\left(1 + \sin\frac{1}{x}\right) - x^2\cos\frac{1}{x} + \frac{2}{x^3}e^{-\frac{1}{x^2}} \qquad \text{if } x \neq 0,$$

$$y' = f'(x) = 0 \qquad \text{if } x = 0,$$

which is continuous for all values of x. But for values of x which satisfy

$$0 < \bar{x} = \frac{1}{2n\pi}$$

we find

$$\cos(1/\bar{x}) = 1, \quad \sin(1/\bar{x}) = 0, \quad f'(\bar{x}) = \bar{x}^2\left(4\bar{x} - 1 + \frac{2}{\bar{x}^5}e^{-\frac{1}{\bar{x}^2}}\right).$$

For sufficiently large values of n, i. e., for sufficiently small values of $\bar{x}$, $f'(\bar{x})$ is evidently negative. Hence for sufficiently large values of n we shall have

$$f'(1/2n\pi) < 0, \quad \text{and likewise} \quad f'(1/(2n+1)\pi) > 0.$$

Hence $f'(x)$ is both positive and negative at points on both sides of $x = 0$ inside of any interval about $x = 0$. And $f'(x)$ is continuous everywhere. *Nevertheless*, $f(x)$ is at an *absolute proper minimum* at $x = 0$.* For we have $f(0) = 0$, and $f(x) > 0$ when $x \neq 0$. Any attempt to establish the fact by any of the usual methods is futile, for $f''(x) = 0$ for $x = 0$ and $f'''(0)$ does not exist, while both f'' and f''' are both positive and negative near $x = 0$ on both sides of that point.

This example depends essentially, as do those which follow, upon the fact established above, viz., that $\xi(x, h)$ does not necessarily take on all values near zero for any limited values of h even though $f'(x)$ is continuous. For if ξ took on all values in the manner specified above, we should have

$$f'(k + t) > 0 \quad \text{and} \quad f'(k - t) < 0 \quad \text{whenever} \quad 0 < t < \delta,$$

since $f(k + h) - f(k) > 0$ and $f(k - h) - f(k) < 0$ provided $x = k$ gives a minimum. In fact we may say that

XX. *If the function* $|\xi(k, h)|$ *takes on all values less than* ${}^2\lambda$ *for some value of* $|h|$ *less than* ${}^1\rho$, *then the condition*

$$f'(k) = 0, \quad f'(k + t) > 0, \quad f'(k - t) < 0, \qquad 0 < t < \delta$$

is both necessary and sufficient for the existence of a proper minimum at $x = k$.

It is evident from the preceding example, however, that $f(x)$ may be at

* That is $f(0) < f(x)$ for all $x \neq 0$.

an absolute proper minimum and that $f'(x)$ may have no settled sign in any interval on either side of the point in question, even though $f'(x)$ is everywhere continuous.* Since the values which ξ takes on for values of h less than any positive number ρ form an assemblage with the power of the continuum, we may at least assert that

XXI. *If a continuous function $f(x)$ is at a proper minimum for $x = k$, the assemblage of values of x for which $f'(x)$ is negative (positive) has the power of the continuum in any interval whatever to the left (right) of $x = k$.*

In a similar manner examples may be constructed in which any desired number of derivatives exist and are continuous, but for which no process except direct inspection of the given function leads to any result; and for which the point in question gives an absolute proper minimum.

It may be objected that in each of the preceding examples the number of maxima and minima is infinite. In the following example:

$$y = \phi(x) = \int_0^x f(x)\,dx,$$

where

$$f(x) = x^3\left(1 + \sin\frac{1}{x}\right) \qquad \text{if } x \neq 0,$$
$$f(x) = 0 \qquad \text{if } x = 0,$$

this objection is removed. Since $f(x)$ is everywhere continuous and has the same sign as x itself, the function $\phi(x)$ is continuous, has its derivative $\phi'(x)\,[=f(x)]$ everywhere continuous, and is at an *isolated absolute proper minimum* for $x = 0$. *Nevertheless* $\phi'(x)$ does not always have the same sign as x, for $f(x)\,[=\phi'(x)]$ vanishes whenever $x = \dfrac{2}{(4n-1)\pi}$, and changes sign at such points.

While the results presented in this paper cannot be said to be complete, they present at least some characteristic properties of the remarkable function $\xi(x, h)$ which occurs in the law of the mean, and the writer hopes that this study may clear up some of the indistinct ideas regarding derivatives which must have troubled many a reader.

COLUMBIA, MISSOURI,
DECEMBER, 1905.

* That is, knowing that $f(x)$, together with $f'(x)$, is everywhere continuous, and knowing that $f(x)$ has a maximum for $x = c$, we can *not* assert that $f'(c + h) \leqq 0$ when $0 < h < \delta$, for example, even though δ be arbitrarily small. This holds even if the maximum is isolated and proper (see next example).